



wwPDB NMR Structure Validation Summary Report ⓘ

Feb 22, 2020 – 01:39 AM EST

PDB ID : 2M5K
EMDB ID : EMD-5590
Title : Atomic-resolution structure of a doublet cross-beta amyloid fibril
Authors : Fitzpatrick, A.W.P.; Debelouchina, G.T.; Bayro, M.J.; Clare, D.K.; Caporini, M.A.; Bajaj, V.S.; Jaroniec, C.P.; Wang, L.; Ladizhansky, V.; Muller, S.; MacPhee, C.E.; Waudby, C.A.; Mott, H.R.; de Simone, A.; Knowles, T.P.J.; Saibil, H.R.; Vendruscolo, M.; Orlova, E.V.; Griffin, R.G.; Dobson, C.M.
Deposited on : 2013-02-27

This is a wwPDB NMR Structure Validation Summary Report for a publicly released PDB/EMDB entry.

We welcome your comments at validation@mail.wwpdb.org

A user guide is available at

<https://www.wwpdb.org/validation/2017/NMRValidationReportHelp>

with specific help available everywhere you see the ⓘ symbol.

The following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

MolProbity : 4.02b-467
Percentile statistics : 20171227.v01 (using entries in the PDB archive December 27th 2017)
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : 2.8

1 Overall quality at a glance

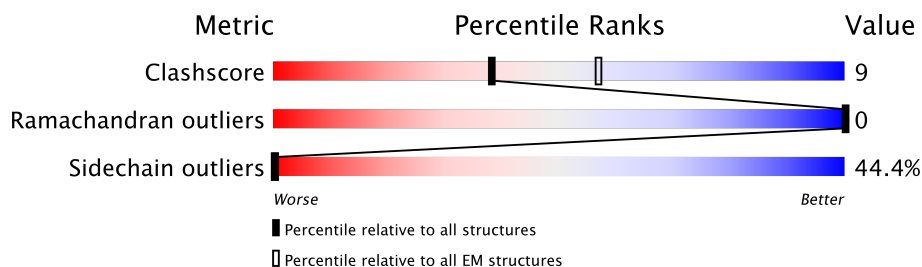
The following experimental techniques were used to determine the structure:

SOLID-STATE NMR, ELECTRON MICROSCOPY

The reported resolution of this entry is 12.70 Å.

The overall completeness of chemical shifts assignment is 47%.

Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.



Metric	Whole archive (#Entries)	NMR archive (#Entries)
Clashscore	136327	1886
Ramachandran outliers	132723	1663
Sidechain outliers	132532	1531

The table below summarises the geometric issues observed across the polymeric chains and their fit to the experimental data. The red, orange, yellow and green segments indicate the fraction of residues that contain outliers for ≥ 3 , 2, 1 and 0 types of geometric quality criteria. A cyan segment indicates the fraction of residues that are not part of the well-defined cores, and a grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions $\leq 5\%$

Mol	Chain	Length	Quality of chain
1	A	11	55% 36% 9%
1	B	11	55% 36% 9%
1	C	11	55% 36% 9%
1	D	11	55% 36% 9%
1	E	11	55% 36% 9%
1	F	11	55% 36% 9%
1	G	11	55% 36% 9%
1	H	11	55% 36% 9%

2 Ensemble composition and analysis ⓘ

This entry contains 1 models. Identification of well-defined residues and clustering analysis are not possible.

3 Entry composition

There is only 1 type of molecule in this entry. The entry contains 1376 atoms, of which 696 are hydrogens and 0 are deuteriums.

In the tables below, the AltConf column contains the number of residues with at least one atom in alternate conformation and the Trace column contains the number of residues modelled with at most 2 atoms.

- Molecule 1 is a protein called Transthyretin.

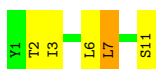
Mol	Chain	Residues	Atoms					AltConf	Trace
1	A	11	Total	C	H	N	O	0	
			172	57	87	11	17		
1	B	11	Total	C	H	N	O	0	
			172	57	87	11	17		
1	C	11	Total	C	H	N	O	0	
			172	57	87	11	17		
1	D	11	Total	C	H	N	O	0	
			172	57	87	11	17		
1	E	11	Total	C	H	N	O	0	
			172	57	87	11	17		
1	F	11	Total	C	H	N	O	0	
			172	57	87	11	17		
1	G	11	Total	C	H	N	O	0	
			172	57	87	11	17		
1	H	11	Total	C	H	N	O	0	
			172	57	87	11	17		

4 Residue-property plots [i](#)

These plots are provided for all protein, RNA and DNA chains in the entry. The first graphic is the same as shown in the summary in section 1 of this report. The second graphic shows the sequence where residues are colour-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. Stretches of 2 or more consecutive residues without any outliers are shown as green connectors. Residues which are classified as ill-defined in the NMR ensemble, are shown in cyan with an underline colour-coded according to the previous scheme. Residues which were present in the experimental sample, but not modelled in the final structure are shown in grey.

• Molecule 1: Transthyretin

Chain A: 



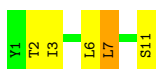
• Molecule 1: Transthyretin

Chain B: 



• Molecule 1: Transthyretin

Chain C: 



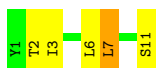
• Molecule 1: Transthyretin

Chain D: 



• Molecule 1: Transthyretin

Chain E: 



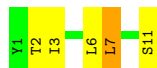
• Molecule 1: Transthyretin

Chain F: 



- Molecule 1: Transthyretin

Chain G:  55% 36% 9%



- Molecule 1: Transthyretin

Chain H:  55% 36% 9%



5 Refinement protocol and experimental data overview

The models were refined using the following method: *simulated annealing*.

Of the 100 calculated structures, 1 were deposited, based on the following criterion: *structures with the least restraint violations*.

The following table shows the software used for structure solution, optimisation and refinement.

Software name	Classification	Version
CNSSOLVE	structure solution	
CNSSOLVE	refinement	

The following table shows chemical shift validation statistics as aggregates over all chemical shift files. Detailed validation can be found in section 7 of this report.

Chemical shift file(s)	2m5k_cs.cif
Number of chemical shift lists	1
Total number of shifts	520
Number of shifts mapped to atoms	520
Number of unparsed shifts	0
Number of shifts with mapping errors	0
Number of shifts with mapping warnings	0
Assignment completeness (well-defined parts)	47%

Note: This is a solid-state NMR structure, where hydrogen atoms are typically not assigned a chemical shift value, which may lead to lower completeness of assignment measure.

No validations of the models with respect to experimental NMR restraints is performed at this time.

6 Model quality

6.1 Standard geometry

There are no covalent bond-length or bond-angle outliers.

There are no bond-length outliers.

There are no bond-angle outliers.

There are no chirality outliers.

There are no planarity outliers.

6.2 Too-close contacts

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in each chain respectively. The H(added) column lists the number of hydrogen atoms added and optimized by MolProbity. The Clashes column lists the number of clashes averaged over the ensemble.

Mol	Chain	Non-H	H(model)	H(added)	Clashes
1	A	85	87	87	3
1	B	85	87	87	3
1	C	85	87	87	3
1	D	85	87	87	3
1	E	85	87	87	3
1	F	85	87	87	3
1	G	85	87	87	3
1	H	85	87	87	3
All	All	680	696	696	12

The all-atom clashscore is defined as the number of clashes found per 1000 atoms (including hydrogen atoms). The all-atom clashscore for this structure is 9.

5 of 12 clashes are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Clash(Å)	Distance(Å)
1:C:7:LEU:HD21	1:D:6:LEU:HD23	0.54	1.79
1:G:7:LEU:HD21	1:H:6:LEU:HD23	0.53	1.79
1:A:7:LEU:HD21	1:B:6:LEU:HD23	0.53	1.80
1:E:7:LEU:HD21	1:F:6:LEU:CD2	0.53	2.34
1:A:7:LEU:HD21	1:B:6:LEU:CD2	0.52	2.34

6.3 Torsion angles

6.3.1 Protein backbone

In the following table, the Percentiles column shows the percent Ramachandran outliers of the chain as a percentile score with respect to all PDB entries followed by that with respect to all NMR entries. The Analysed column shows the number of residues for which the backbone conformation was analysed and the total number of residues.

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	A	9/11 (82%)	9 (100%)	0 (0%)	0 (0%)	100	100
1	B	9/11 (82%)	9 (100%)	0 (0%)	0 (0%)	100	100
1	C	9/11 (82%)	9 (100%)	0 (0%)	0 (0%)	100	100
1	D	9/11 (82%)	9 (100%)	0 (0%)	0 (0%)	100	100
1	E	9/11 (82%)	9 (100%)	0 (0%)	0 (0%)	100	100
1	F	9/11 (82%)	9 (100%)	0 (0%)	0 (0%)	100	100
1	G	9/11 (82%)	9 (100%)	0 (0%)	0 (0%)	100	100
1	H	9/11 (82%)	9 (100%)	0 (0%)	0 (0%)	100	100
All	All	72/88 (82%)	72 (100%)	0 (0%)	0 (0%)	100	100

There are no Ramachandran outliers.

6.3.2 Protein sidechains

In the following table, the Percentiles column shows the percent sidechain outliers of the chain as a percentile score with respect to all PDB entries followed by that with respect to all NMR entries. The Analysed column shows the number of residues for which the sidechain conformation was analysed and the total number of residues.

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	A	9/9 (100%)	5 (56%)	4 (44%)	0	2
1	B	9/9 (100%)	5 (56%)	4 (44%)	0	2
1	C	9/9 (100%)	5 (56%)	4 (44%)	0	2
1	D	9/9 (100%)	5 (56%)	4 (44%)	0	2
1	E	9/9 (100%)	5 (56%)	4 (44%)	0	2
1	F	9/9 (100%)	5 (56%)	4 (44%)	0	2
1	G	9/9 (100%)	5 (56%)	4 (44%)	0	2
1	H	9/9 (100%)	5 (56%)	4 (44%)	0	2

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
All	All	72/72 (100%)	40 (56%)	32 (44%)	0 2

5 of 32 residues with a non-rotameric sidechain are listed below. They are sorted by the frequency of occurrence in the ensemble.

Mol	Chain	Res	Type
1	D	7	LEU
1	G	11	SER
1	A	2	THR
1	C	11	SER
1	F	6	LEU

6.3.3 RNA [i](#)

There are no RNA molecules in this entry.

6.4 Non-standard residues in protein, DNA, RNA chains [i](#)

There are no non-standard protein/DNA/RNA residues in this entry.

6.5 Carbohydrates [i](#)

There are no carbohydrates in this entry.

6.6 Ligand geometry [i](#)

There are no ligands in this entry.

6.7 Other polymers [i](#)

There are no such molecules in this entry.

6.8 Polymer linkage issues [i](#)

There are no chain breaks in this entry.

7 Chemical shift validation

The completeness of assignment taking into account all chemical shift lists is 47% for the well-defined parts and 47% for the entire structure.

7.1 Chemical shift list 1

File name: 2m5k_cs.cif

Chemical shift list name: *assigned_chem_shift_list_1*

7.1.1 Bookkeeping

The following table shows the results of parsing the chemical shift list and reports the number of nuclei with statistically unusual chemical shifts.

Total number of shifts	520
Number of shifts mapped to atoms	520
Number of unparsed shifts	0
Number of shifts with mapping errors	0
Number of shifts with mapping warnings	0
Number of shift outliers (ShiftChecker)	0

7.1.2 Chemical shift referencing

The following table shows the suggested chemical shift referencing corrections.

Nucleus	# values	Correction \pm precision, ppm	Suggested action
$^{13}\text{C}_\alpha$	88	0.42 ± 0.13	None needed (< 0.5 ppm)
$^{13}\text{C}_\beta$	80	-1.25 ± 0.07	Should be applied
$^{13}\text{C}'$	88	1.44 ± 0.10	Should be applied
^{15}N	72	-2.13 ± 0.28	Should be applied

7.1.3 Completeness of resonance assignments

The following table shows the completeness of the chemical shift assignments for the well-defined regions of the structure. The overall completeness is 47%, i.e. 472 atoms were assigned a chemical shift out of a possible 1000. 16 out of 16 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	^1H	^{13}C	^{15}N
Backbone	240/424 (57%)	0/168 (0%)	176/176 (100%)	64/80 (80%)
Sidechain	168/448 (38%)	0/264 (0%)	168/184 (91%)	0/0 (—%)

Continued on next page...

Continued from previous page...

	Total	¹H	¹³C	¹⁵N
Aromatic	64/128 (50%)	0/64 (0%)	64/64 (100%)	0/0 (—%)
Overall	472/1000 (47%)	0/496 (0%)	408/424 (96%)	64/80 (80%)

Note: This is a solid-state NMR structure, where hydrogen atoms are typically not assigned a chemical shift value, which may lead to lower completeness of assignment measure.

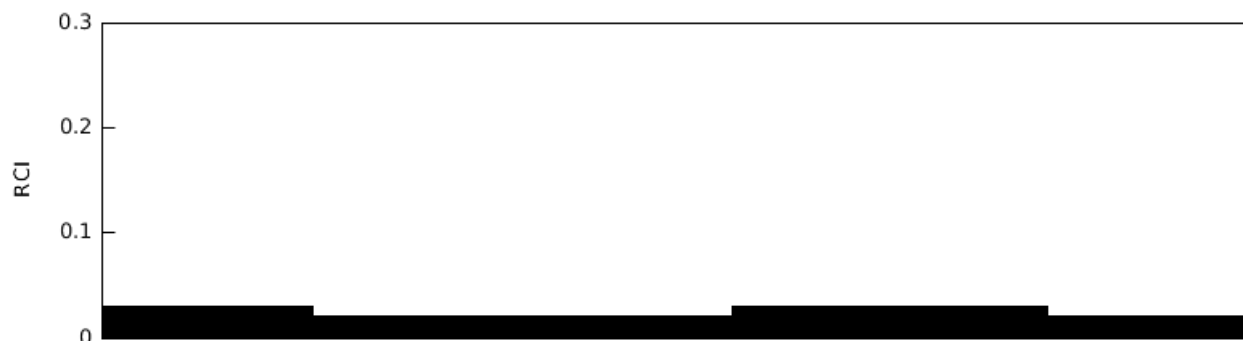
7.1.4 Statistically unusual chemical shifts [i](#)

There are no statistically unusual chemical shifts.

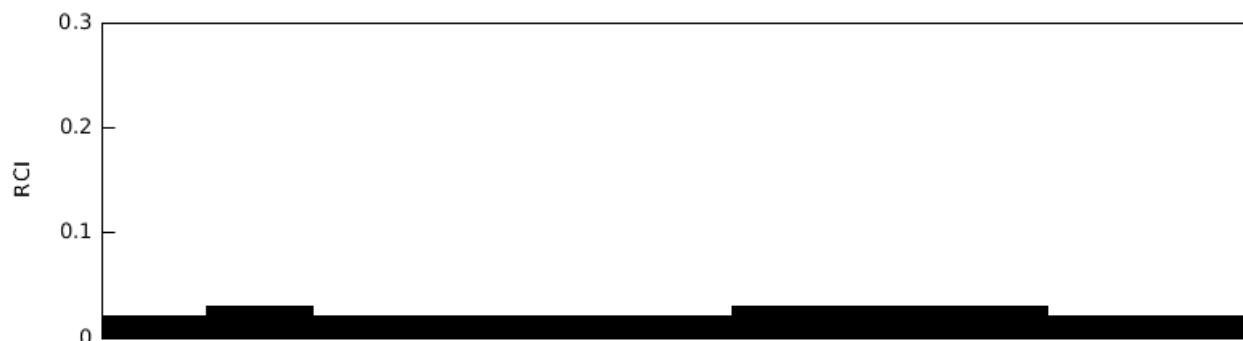
7.1.5 Random Coil Index (RCI) plots [i](#)

The images below report *random coil index* values for the protein chains in the structure. The height of each bar gives a probability of a given residue to be disordered, as predicted from the available chemical shifts and the amino acid sequence. A value above 0.2 is an indication of significant predicted disorder. The colour of the bar shows whether the residue is in the well-defined core (black) or in the ill-defined residue ranges (cyan), as described in section 2 on ensemble composition.

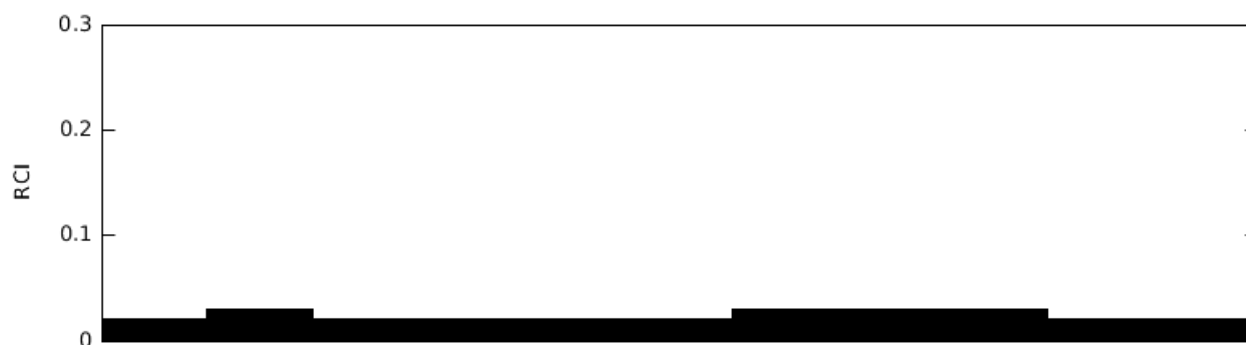
Random coil index (RCI) for chain A:



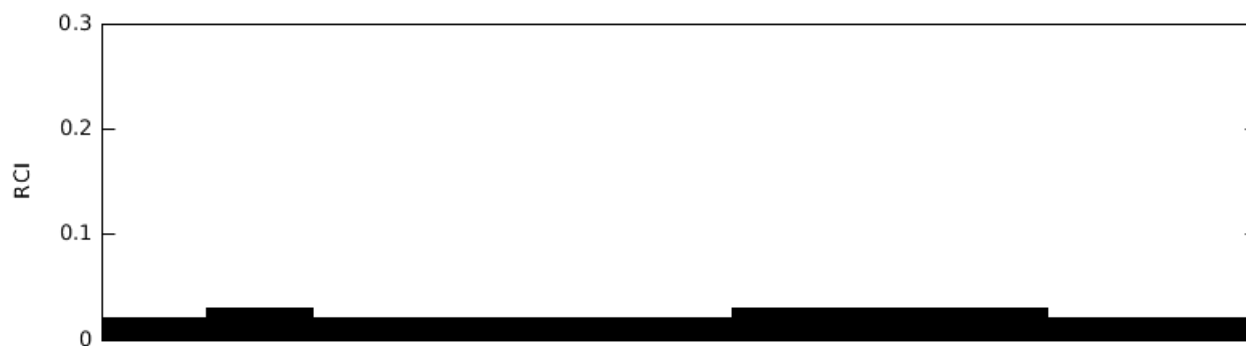
Random coil index (RCI) for chain B:



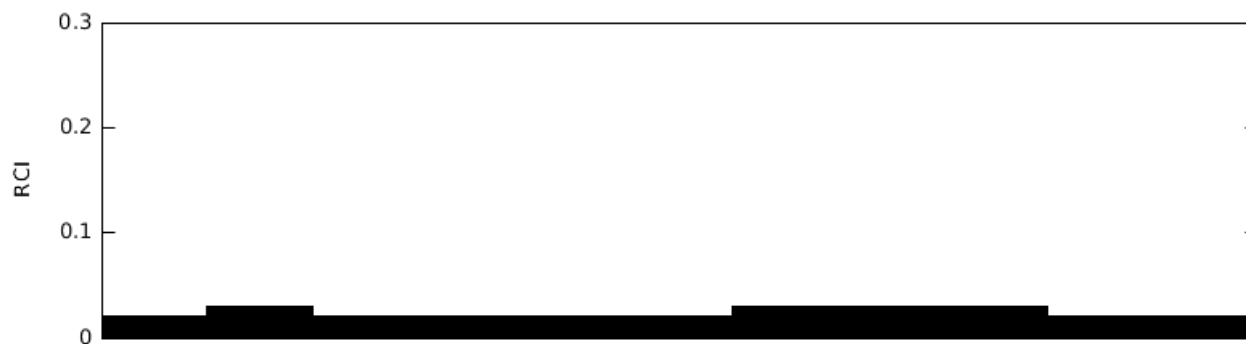
Random coil index (RCI) for chain C:



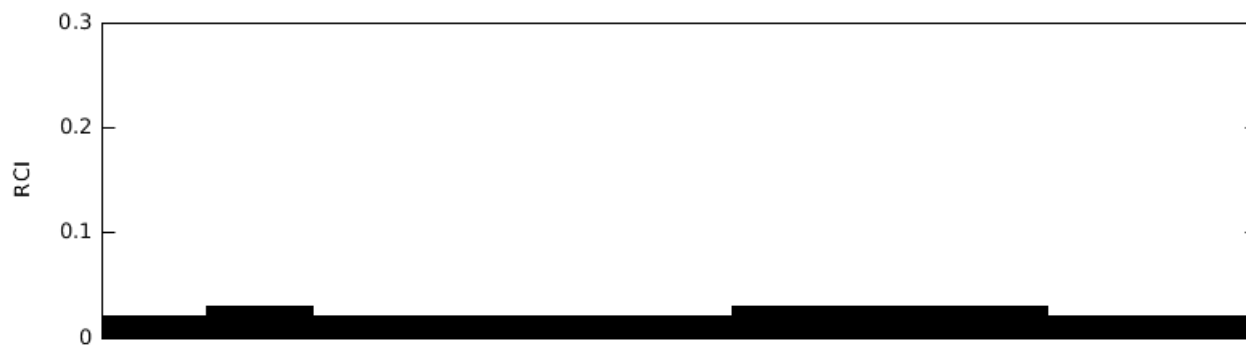
Random coil index (RCI) for chain D:



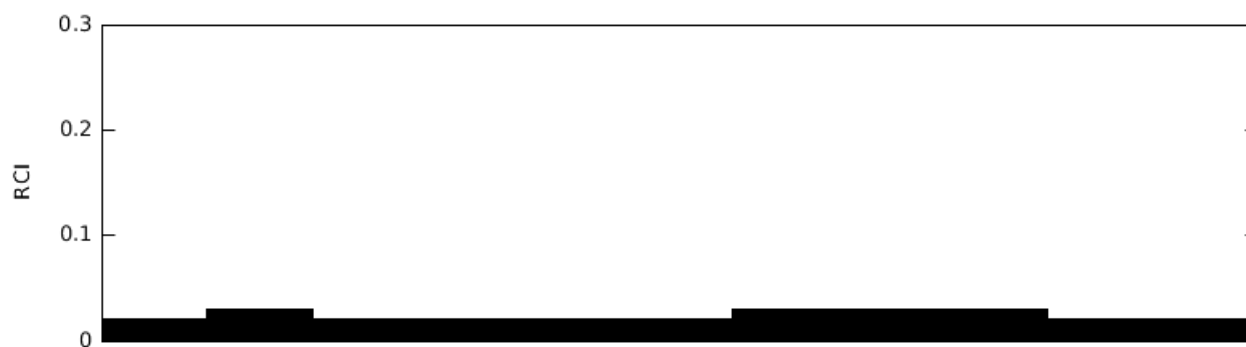
Random coil index (RCI) for chain E:



Random coil index (RCI) for chain F:



Random coil index (RCI) for chain G:



Random coil index (RCI) for chain H:

